

# Robot George - Interactive Continuous Learning of Visual Concepts

Michael Zillich, Kai Zhou  
Vienna University of Technology, Austria  
{zillich,zhou}@acin.tuwien.ac.at

Danijel Skočaj, Matej Kristan, Alen Vrečko, Marko Mahnič  
University of Ljubljana, Slovenia  
{danijel.skocaj,matej.kristan,alen.vrecko,marko.mahnic}@fri.uni-lj.si

Miroslav Janíček, Geert-Jan M. Kruijff  
DFKI, Saarbrücken, Germany  
{miroslav.janicek,gj}@dfki.de

Thomas Keller  
Albert-Ludwigs-Univ. Freiburg, Germany  
tkeller@informatik.uni-freiburg.de

Marc Hanheide, Nick Hawes  
University of Birmingham, UK  
{m.hanheide,N.A.Hawes}@cs.bham.ac.uk

**Abstract**—The video presents the robot George learning visual concepts in dialogue with a tutor.

**Keywords**—Robotics; Algorithms; Human Robot Interaction; Learning; Concepts

## I. INTRODUCTION

Interactive continuous learning is an important characteristic of a cognitive agent that is supposed to operate and evolve in an everchanging environment. In the video we present an integrated system for learning visual concepts in dialogue with a tutor. We designed and implemented the representations and mechanisms that are needed in such a continuous learning process. We implemented an approach for modelling beliefs that are created by processing visual and linguistic information. Based on the detected knowledge gaps represented in the beliefs, the motivation and planning mechanism implements four types of interaction for learning. The implemented mechanisms are briefly outlined in this document and references to the papers that contain more in-depth descriptions are also given.

## II. THE SYSTEM

*Visual processing* serves to provide the object hypotheses together with visual properties about which the system will subsequently learn. Given that the system learns from a variety of as yet unknown objects, we implemented a generic active object detection and segmentation scheme that uses 2D and 3D information, exploiting the fact that objects are presented on planar supporting surfaces [1].

*The models of visual properties* like colour and shape are represented as generative models based on multivariate online discriminative Kernel Density Estimator (odKDE) [2] and are constructed in an online fashion from new observations. Similarly, also the object models are learned online when the new objects are presented to the robot [3].

*Situated dialogue understanding* and production is treated as an abductive problem. Language understanding is treated as inference to the most appropriate intention and beliefs behind a communicative act, whereas production is inference to the most appropriate realization of the robot's (communicative) intention and beliefs. Given a goal, the abductive reasoner builds up and continually refines a set of partial defeasible explanations of the input, conditioned on the verification of the knowledge gaps they contain [4].

*Beliefs* express factual information about the state of the world. In our approach, they are relational structures that account for the inherent uncertainty using multivariate probability distributions over properties and their values. They are situated, anchored to a given situation, and mutually interlinked.

*The motivation management* [5] monitors the beliefs and based on them creates goals and selects which of them to pass on to planning. *The planner* [6] then builds a plan to satisfy a given goal, which is subsequently executed. In this way a system behaviour is generated and controlled.

*The system architecture* is based on CAS (CoSy Architecture Schema) [7], which is essentially a distributed working memory model. Using this architecture, a complex, distributed, asynchronous, and heterogeneous *system* has been built [8] and is presented in the video.

## ACKNOWLEDGMENT

This work was supported by the EU FP7 ICT cognitive systems integrated project CogX-215181.

## REFERENCES

- [1] K. Zhou, A. Richtsfeld, M. Zillich, M. Vincze, A. Vrečko, and D. Skočaj, "Visual information abstraction for interactive robot learning," in *ICAR 2011*, Tallinn, Estonia, June 2011.
- [2] M. Kristan and A. Leonardis, "Online discriminative kernel density estimation," in *ICPR 2010*, Istanbul, Turkey, August 2010, pp. 581–584.
- [3] M. Zillich, J. Prankl, T. Mörwald, and M. Vincze, "Knowing your limits - self-evaluation and prediction in object recognition," in *IROS 2011*, San Francisco, CA, USA, 2011.
- [4] M. Janíček, "Abductive reasoning for continual dialogue understanding," in *Proceedings of the ESSLLI 2011*, Ljubljana, Slovenia, 2011.
- [5] M. Hanheide, N. Hawes, J. Wyatt, M. Göbelbecker, M. Brenner, K. Sjöö, A. Aydemir, P. Jensfelt, H. Zender, and G.-J. Kruijff, "A framework for goal generation and management," in *Proceedings of the AAAI Workshop on Goal-Directed Autonomy*, 2010.
- [6] M. Brenner and B. Nebel, "Continual planning and acting in dynamic multiagent environments," *Journal of Autonomous Agents and Multiagent Systems*, vol. 19, no. 3, pp. 297–331, 2009.
- [7] N. Hawes and J. Wyatt, "Engineering intelligent information-processing systems with CAST," *Adv. Eng. Inform.*, vol. 24, no. 1, pp. 27–39, 2010.
- [8] D. Skočaj, M. Kristan, A. Vrečko, M. Mahnič, M. Janíček, G.-J. M. Kruijff, M. Hanheide, N. Hawes, T. Keller, M. Zillich, and K. Zhou, "A system for interactive learning in dialogue with a tutor," in *IROS 2011*, San Francisco, CA, USA, 25-30 September 2011, pp. 3387–3394.